Reprint from

Current Topics in Microbiology and Immunology, Vol. 240
Plant Biotechnology: New Products and Applications
Editors: J. Hammond, P. McGarvey, and V. Yusibov

Springer

Transgenic Plants for Therapeutic Proteins: Linking Upstream and Downstream Strategies

C.L. Cramer[1], J.O. Boothe[2], and K.K. Oishi[3]

1 Introduction

With the new knowledge generated through the Human Genome Project and related biomedical research comes a potential revolution in drug development strategies. One of the most direct applications of this knowledge will be highly specialized recombinant protein-based therapeutics. Recombinant drugs such as human erythropoietin (EPO), tissue plasminogen activator (tPA), and Cerezyme® (glucocerebrosidase) are currently on the market and many other recombinant proteins are in various stages of human clinical trials. Commercial production of

[1] CropTech Corp., Virginia Tech Corporate Research Center, Blacksburg, VA, 24060 and Department of Plant Pathology, Physiology and Weed Science, Fralin Biotechnology Center, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0346, USA
[2] Sam Rudy's Cosmetics Inc., Suite 304, 609 14th St. N.W., Calgary, Alberta, T2N, Canada
[3] CropTech Corp., Virginia Tech Corporate Research Center, Blacksburg, VA, 24060, USA

these proteins utilizes fermentation (primarily *E. coli* and yeast) and mammalian cell systems (e.g., Chinese hamster ovary cells), the major expression systems adopted by the well established biotechnology companies. However, these expression systems have significant limitations. Bacteria cannot perform the complex posttranslational modifications required for bioactivity of many human proteins and high-level expression often leads to accumulation of insoluble protein aggregates. While mammalian cell cultures perform the required protein modifications, low transgene expression levels, instability of selected cell lines, and the difficulties and high expense of scale-up are often limiting or severely impact cost. Thus, there remains significant opportunity for alternative expression systems that address these limitations and cost issues to compete in the protein therapeutics market. In fact, development of more cost-effective protein bioproduction systems may be critical in translating the discoveries of genomics and medical research into widely available and affordable treatments and cures. Recent advances in the area of transgenics – the use of genetically engineered plants and animals for bioproduction – indicate great promise as effective protein factories. The fact that recombinant proteins from both transgenic animals and transgenic plants are now in clinical trials demonstrates significant progress toward commercialization of these technologies.

For any particular target protein, selection of a recombinant system will depend on the characteristics of the desired protein product, the volume needs (size of the market), and market-driven cost constraints (reviewed by Pen 1996). Transgenic plants have some remarkable features that make them particularly well suited for cost-effective bioproduction of proteins for pharmaceutical uses. These include; (a) low production costs, (b) reduced time to market, (c) unlimited supply, (d) eukaryotic protein processing, and (e) safety. Cost advantages are based not only on the low cost of biomass production, but also costs associated with research and development, germplasm scale-up (e.g., imagine the infrastructure investment of tripling the capacity of one's aseptic fermentation or mammalian cell production facility compared to tripling one's acreage for plant growth), and reduced requirements for quality assurance testing for exclusion of human pathogenic agents (reviewed in Owen and Pen 1996). Plant-based strategies also have advantages in the pace at which feasibility testing can be done and R & D successes can be scaled up and brought to market. For example, a tobacco plant goes from seed to next generation seed in three months and produces up to a million seed per plant. Scaling-up to hundred or thousands of acres is very rapid.

Many of the therapeutic proteins of interest require complex posttranslational processing and/or oligomerization for bioactivity or appropriate targeting following administration to patients. There appears to be remarkable conservation of these protein processing steps between plants and animals such that the majority of human proteins that have been produced in plants (see Table 1) show significant structural, biochemical and functional equivalency to proteins from humans or animal cell cultures. In cases where certain modification steps are lacking or differ in plants (e.g., glycan composition, discussed further below), strategies to introduce appropriate animal protein processing enzymes or modify the plant processing

Table 1. Fidelity of plant-based production of human (or other animal) proteins

Transgene product	Potential use disease target	Plant host	Structural integrity	Functional activity	Reference
Serum proteins					
Hemoglobin (α and β)	Blood substitute	Tobacco	Yes (dimer)	Yes (O_2/CO_2 binding)	Dyrnyck et al. 1997
Human serum albumin	Blood extender	Potato	Yes	Not tested	Sijmons et al. 1990
Protein C	Anticoagulant	Tobacco	Most processing steps performed	Not tested	Cramer et al. 1996a
Cytokines/lymphokines					
α-Interferon	Viral protection, anti-cancer	Rice	Yes	Yes (viral resistance assay)	Zhu et al. 1994
γ-Interferon	Phagocyte activator	Tobacco	Yes (glycans[a])	Yes (in vitro assay)	Ghill 1997
GM-CSF	Leukopenia in bone marrow transplants	Tobacco	Yes (glycans[a])	Yes (growth stimul. of TF-1 cells)	Giddi et al. 1996
Epidermal growth factor	Mitogen	Tobacco	CRIM[b]	Not tested	Higo et al. 1995
Tissue growth factor	Mitogen	Tobacco	Yes (glycans[a])	Not tested	Biocca et al. 1994
Erythropoietin	Mitogen, blood cells	Tobacco cells	Yes (glycans[a])	Not tested	Matsumoto et al. 1995
Lysosomal enzymes					
α-Galactosidase	Fabry disease	Tobacco	Yes (glycans[a])	Yes (enzyme act.)	Ganz, 1997
Glucocerebrosidase	Gaucher disease	Tobacco	Yes (glycans[a])	Yes (enzyme act.)	Cramer et al. 1996b
Other proteins					
Casein	Nutriceutical	Potato	CRIM[b]	Not tested	Chong et al. 1997
Hirudin	Anticoagulant	Canola	Yes	Yes (thrombin inh.)	Parmenter et al. 1995
NP1 defensin	Antibiotic	Tobacco	CRIM[b]	Yes (antibiotic act.)	Owen 1997
Glutamate decarboxylase	Diabetes	Tobacco	CRIM[b]	Yes (mouse model)	Ma et al. 1997

[a]Proteins were glycosylated but the glycan composition may differ from those produced in humans
[b]Detected as cross-reactive immunodetected material by western immunoblots or ELISA

machinery are greatly facilitated by the ease of plant transformation and the broad experience in transgenic approaches to modifying plant metabolism through overexpression and antisense strategies. In fact, plants may be the only system capable of efficient production of certain human proteins such as growth regulators and cell cycle inhibitors which would negatively impact either the transgenic animal or animal cell culture in which they are expressed.

Perhaps the most important advantage of plants, which is emerging in the aftermath of the recent "mad cow disease" scare, involves product safety. The biopharmaceutical industry is now faced with the possibility of product validation

and quality assurances that demonstrate purity not only from known human pathogens such as HIV but also from unknown or poorly characterized agents such as the prions responsible for bovine spongiform encephalopathy and the related Creutzfeld-Jakob disease (Roberts and James 1996; Vaughan 1996). Plants do not serve as hosts for blood- or animal tissue-borne human pathogens. In addition, plant-based production and purification can be executed without the use of any animal-derived products. Clearly purity, efficacy and quality control issues similar to production of any biopharmaceutical will need to be addressed (see Miele 1997). However, plant-based bioproduction should realize substantial savings as a human- and animal-source-free production system.

The list of complex human proteins and animal, viral and bacterial proteins of medical value that have been successfully expressed in plants is growing rapidly (reviewed in Owen and Pen 1996). In addition to disease antigens (vaccines) and antibodies discussed in other chapters of this volume, transgenic plants have been used to synthesize a number of complex serum proteins, cytokines, growth regulators, anticoagulants, antibiotics, and lysosomal enzymes (see Table 1). Most of these proteins appear fully functional and structurally comparable to the analogous proteins produced in animal cell cultures or in humans. Thus, plants have clearly passed the initial test of feasibility – they are capable of producing bioactive human proteins of pharmaceutical value. In addition, the first transgenic plant-synthesized products (a tobacco-derived antibody targeting gum disease and a potato-derived edible vaccine candidate) have reached initial human trials – a significant benchmark toward commercialization. However, as we move from feasibility studies to commercial bioproduction, issues of transgene expression levels, product processing and stability, biomass and extraction scale-up, purification, and quality control become paramount. These longer-term goals have inspired the development of novel transgene expression systems that incorporate components targeting product abundance, product recovery, and regulatory acceptance into the initial transgene design. In this review, we will discuss key issues that impact the choice and utility of plant-based production systems for biopharmaceuticals. We will highlight several production strategies that stress the importance of linking "upstream" steps in genetic engineering and expression strategies with "downstream" issues of extraction, purification, and yield. These systems are designed to separate biomass production from transgenic protein production and to directly manipulate the timing, tissue and subcellular localization of the product to enhance yield, protein stability, and ease of recovery and purification.

2 Plant-Based Biopharmaceutical Production: Issues and Answers

The majority of examples demonstrating bioproduction of potential therapeutic proteins in plants shown in Table 2 have used model plant species that are easy to genetically engineer (e.g., tobacco, potato) and the "strong, constitutive" 35S

Table 1. Transgene expression strategies and recombinant protein yield

Transgene product	Plant host	Promoter	Expression strategy, tissue	Production levels (% of soluble protein)	Reference
Serum proteins					
Hemoglobin (α and β)	Tobacco	35S[a]	Seed, root	0.05%	Dieryck et al. 1997
Human serum albumin	Potato	35S[a]	Constitutive, leaf	0.02%	Sijmons et al. 1990
Protein C	Tobacco	35S	Constitutive, leaf	0.002%	Cramer et al. 1996a
Cytokines/lymphokines					
α-Interferon	Rice	PI[a]	Constitutive, leaf	Not reported	Zhu et al. 1994
γ-Interferon	Tobacco	NA	Genewave viral-infected leaf	1%	Gruel 1997
GM-CSF	Tobacco	Rice glutelin	Seed-specific	Not reported	Geier et al. 1996
Epidermal growth factor	Tobacco	35S	Constitutive, leaf	0.001%	Higo et al. 1993
Lysosomal enzymes					
α-Galactosidase	Tobacco	NA	Genewave/leaf	12.1mg/kg tissue	Gruel 1997
Glucocerebrosidase	Tobacco	MeGA	Post-harvest infused leaf tissues	1%–10%	Cramer et al. 1996b
Viral or bacterial antigens					
E. coli enterotoxin B	Potato	35S[a]	Microtuber	0.003%	Haq et al. 1995
Cholera toxin	Tobacco	35S[a]	Constitutive, leaf	Not reported	Haq et al. 1995
Hepatitis B surf. antigen	Tobacco	35S[a]	Constitutive, leaf	0.007%	Mason et al. 1992
Other proteins					
Hirudin	Canola	Oleosin	Seed-specific	0.3% (seed protein)	Parmenter et al. 1995
Antibodies	Tobacco	35S	Constitutive	0.01%–5%	Rev. in Owen and Pen 1996
α-Trichosantin	Tobacco	NA	Genewave/leaf	4%–5%	Kumagai et al. 1993
Glutamate decarboxylase	Tobacco	35S[a]	Constitutive, leaf	0.4%	Ma et al. 1997
	Potato	35S[a]	Tuber	0.4%	Ma et al. 1997

[a]Modified 35S promoter containing enhancer duplication and/or leader sequences (translational enhancer) from the tobacco etch virus or alfalfa mosaic virus

promoter derived from the cauliflower mosaic virus. However, as plant biotechnology moves from demonstrating feasibility toward commercialization of protein products, many other issues come into play in selecting host species, expression strategies, target tissues, and extraction/purification protocols. These choices must take into account not only the production of the particular protein of interest but

issues of recovery, purity, production/purification costs, reproducibility, supply continuity, quality control, and regulatory assessment.

2.1 Selection of Crop Species

While certain features such as low production costs and high biomass capacity are common to all plant-based expression systems, other factors may strongly influence the choice of one plant species or expression strategy over another for the production of a specific foreign protein. In selecting a particular species it is important to consider how readily it can be manipulated to produce a stable transgenic line, the tissue and subcellular compartment best suited for stable expression of the heterologous protein, and the availability of methods for the efficient harvesting and initial processing of the plant material. Included in the first consideration are factors such as the amenability to transformation and regeneration of whole plants, generation time, and tractability to controlled genetic crossing. All of these factors significantly impact upon the time and resources required for product development. Plant transformation technologies are highlighted in other chapters (Hansen and Chilton and Finer et al., this volume) and have been recently reviewed (Lindsey 1996) and are therefore not discussed in detail here. The remaining two considerations deal mainly with product biocomparability (bioactivity, conformation, efficacy) and recovery. Because infrastructure and methods for the harvest and processing of the major crop species already exist, whenever possible these are the species of choice. The tissue and subcellular compartment of expression determines protein processing capabilities, stability of the product and the ease with which it can be recovered.

Tobacco remains the easiest plant to genetically engineer and is widely used to test suitability of plant-based systems for bioproduction of recombinant proteins (see Table 2). Although tobacco is considered a regional crop and relatively labor intensive, at least three plant-based biotech companies are targeting tobacco for biopharmaceutical production (CropTech Corp; BioSource Technologies, Inc. and Planet Biotechnology). In addition to being easily engineered, tobacco is an excellent biomass producer (in excess of 40 tons leaf fresh weight/acre based on multiple mowings per season) and prolific seed producer (up to one million seeds produced per plant), thus hastening the time in which a product can be scaled up and brought to market.

Several companies are developing production strategies involving transgene product accumulation in seeds, an organ designed to accumulate and store protein reserves (see Sect. 2.2). Companies targeting seed-based production using canola, corn or soybeans include Sem BioSys Genetics, Agracetus (USA), Mogen International (the Netherlands), and Plantzyme (the Netherlands). Applied Phytologics (API, Davis, CA) is using transgenic rice and barley seed but is producing and recovering recombinant proteins during seed germination in a process analogous to malting. Other crops being developed for biopharmaceutical protein or vaccine production include alfalfa, banana, potato, and tomato.

2.2 Choice of Tissue

In order to obtain maximum yields, the plant species selected must concentrate biomass in the organ or tissue where the foreign protein is expressed. The diversity among different species in this respect means that a variety of options are available including leaves, vegetative storage organs (e.g. tubers) and seeds. The tissue chosen should be compatible with the desired protein, enabling correct processing, stable accumulation and, whenever possible, efficient recovery. Many human therapeutic proteins require extensive processing for full activity, involving transport through the cellular endomembrane system. Functional lysosomal enzymes (Cramer et al. 1996b) and mammalian antibodies (Ma et al. 1995) have all been produced in leaves of tobacco following trafficking through the endoplasmic reticulum (ER) and Golgi complex. Human serum albumin has also been stably expressed in tobacco leaves and various tissues of potato including tubers (Sijmons et al. 1990), although the precise folding and functionality of the protein was not established. In the above examples the recombinant proteins were either specifically targeted to, and detected in the apoplast, or presumed to locate there as a result of the default pathway of the plant endomembrane system. Deposition into the extracellular apoplast may contribute to the stability of foreign proteins by removing them from the more hydrolytic intracellular environment (Fiers et al. 1993).

Expression and accumulation of foreign proteins in seeds may be achieved through compartmentalization within various subcellular storage organelles. As a natural storage organ, seeds possess attributes such as high protein content and a low hydrolytic intracellular environment that make them particularly attractive as protein production vehicles. The human neuropeptide, leu-enkephalin (Vandekerckhove et al. 1989), and the leech anticoagulant protein, hirudin (Parmenter et al. 1995), have both been produced in seeds of *Brassica napus* following targeting to protein bodies and oil bodies respectively. Proteins can also be secreted to the apoplast of seeds. However, the recovery of apoplastic proteins from seeds may be more difficult than from some of the vegetative organs mentioned above, owing to the desiccated state of seed tissue at maturity. On the other hand, this advanced state of dehydration also confers enhanced stability, allowing seeds to be stored for periods of several years without any appreciable degradation of proteins or loss of activity (e.g., see Pen et al. 1993). The greater flexibility resulting from the separation of protein production and purification represents a distinct advantage of seeds over most other organs for which more immediate processing is often required.

2.3 Expression Strategies

Choice of promoter, which mediates the timing, tissue-specificity, and level of transgene expression, is a key determinant of transgene product yields and recovery strategies (see review by Chinn and Comai 1996). As shown in Table 2, many of the human (or other animal) proteins expressed in plants have used native or enhanced

versions of the 35S promoter derived from the cauliflower mosaic virus to drive "constitutive" transgene expression, and it remains the most widely used promoter in plant biology for over-expression of plant proteins or inhibition via antisense strategies. The 35S promoter is active in most plant tissues (Benfey et al. 1989; Fang et al. 1989) and especially in its modified forms (Kay et al. 1987; Carrington and Freed 1990) can drive quite high levels of protein production. Although most of the human proteins produced using the 35S promoter (Table 2) showed accumulation levels below 0.1% of soluble protein, several transgene products (chitinase, antibodies) have been expressed at levels of 2%–5% of extractable protein. The 35S promoter is quite active during seed development and has been used in production systems targeting recovery of recombinant proteins from seed. However, the 35S promoter (and constitutive expression in general) has significant limitations when commercial bioproduction in nonseed tissues is the goal. Proteins that accumulate to high levels may negatively impact yield or overall health of the plant. High constitutive expression is sometimes associated with co-suppression or gene silencing (Taylor 1997) resulting in little or no transgene product accumulation. For proteins that are not highly stable, constitutive expression can lead to wasteful synthesis-degradation cycles and, of particular detriment for pharmaceutical application, contamination of the final product with inactive degradation products. In addition, the 35S is not highly active in many mature tissues (e.g., mature roots and fully expanded leaves) so that the full potential of biomass cannot be utilized. Use of inducible promoters or promoters that have a tight pattern of tissue- or organ-specificity avoids many of these limitations and appears to be the strategy of choice for most companies targeting plant-based production of high-value proteins.

CropTech scientists have developed a postharvest expression system that uses an inducible promoter termed the MeGA™ promoter (Cramer and Wissemann 1997). This promoter has been modified from a defense-related gene such that it is generally inactive during normal growth and development but shows rapid and strong gene activation in response to mechanical stress (wound-induction, or mechanical gene activation) or a variety of defense elicitors. Thus, the recombinant protein is not synthesized in tobacco leaves in the field (or greenhouse). Plants can be harvested and stored for weeks in a cold room. Recombinant protein production is then induced *de novo* in the laboratory or GMP facility and newly synthesized protein recovered 8–24h later. Because survival depends on both the speed and intensity with which a plant can activate its defenses, we find the MeGA™ promoter highly effective in driving high levels of inducible expression in all tissues of the plant including fully expanded leaves. The postharvest expression strategy has several advantages for pharmaceutical production. Biomass production is both temporally and spatially separated from recombinant product production minimizing the impact of (a) environmental factors on protein yield and quality and (b) possible deleterious effects of transgene expression or foreign protein accumulation on plant growth and development. All recovered protein is newly synthesized. In addition, the timing of protein extraction can be adjusted based on the stability of the particular gene product to optimize yield of fully active polypeptides. For products requiring activation of multiple genes (e.g., multiple subunits, or target proteins that require specialized protein-modifying enzymes), coinduction assures coordinated synthesis. In theory, the postharvest system could also permit further manipulation of the protein synthesis and processing machinery through addition of specific chemicals to the induction medium (e.g., inhibitors of key protein modification steps), although this could add significant expense to commercial scale bioproduction.

Bioproduction strategies involving developmentally defined- or virally vectored-expression (e.g., Biosource's Geneware system) are also designed to limit recombinant protein production to a discrete period. With the Geneware system, TMV-susceptible tobacco is field grown to an appropriate age, inoculated with genetically modified virus, and harvested 2-3 weeks later for recombinant protein extraction (Ganz, 1997). Within this period, the virus titers reach high levels leading to significant transgene product accumulation. Using this system, Biosource scientists have attained very high product yield (recombinant protein representing greater than 10% of total soluble protein) and have progressed to the point of large scale field production and pilot plant extraction. Applied Phytologics utilizes a germination-specific promoter to direct transgene expression. Recombinant protein is produced under controlled conditions following inhibition and initiation of germination of transgenic seed, a production scheme analogous to barley malting. Expression strategies involving seed-based accumulation of recombinant proteins also take advantage of discrete bioproduction periods and separation of transgene activity from the bulk of plant growth. A large number of seed-specific promoters, often derived from genes encoding seed storage proteins, are available for both monocot and dicot plants. Depending on the recovery strategy (see below) and the characteristics of the protein product, promoters specific for embryo- versus endosperm-specific expression can be selected.

2.4 Posttranslational Processing

In comparison with industrial enzyme production, bioproduction of human proteins for pharmaceutical applications is particularly challenging due to the rigorous requirements with respect to purity, reproducibility, efficacy, and biocomparability. Many of the proteins with greatest promise as therapeutics require complex post-translational modifications and/or assembly. The striking fidelity with which plants appear to recognize and correctly act upon most of the processing signals encrypted within mammalian polypeptides indicates a high degree of conservation in protein processing machinery between plants and animals. Conserved processes include endomembrane targeting, signal peptide cleavage, protein folding and oligomerization, disulfide bond formation (although precise cysteine-cysteine bonding patterns have not been directly determined), asparagine-linked glycosylation, selective retention in the ER and Golgi, and C-terminal isoprenylation. We have also noted internal proteolytic processing events in several human proteins expressed in tobacco that appear to mimic processing that occurs in mammalian cells

although the precise termini of the products have not yet been determined (Oishi et al. unpublished data).

However, clear differences in protein processing, most notably in glycoprotein processing, do exist between plants and animals. The glycan moiety of mammalian glycoproteins functions in protein folding and assembly, subcellular targeting, cell- or tissue-specific delivery within the body, protein half-life, and clearance from the bloodstream (Varki 1993). Thus, changes in glycan composition or arrangement are likely to affect activity or pharmacokinetics (Jenkins et al. 1996; Lis 1992). Plant N-linked glycans do not contain terminal sialic acid residues or mannose-6-phosphate and contain other sugars or sugar linkages not found in mammalian glycoproteins. The same amino acid signature (N-X-S/T) is recognized within the ER for addition of the high-mannose form glycan complex (identical in plant and animals). However, plants process these N-linked glycans to distinct complex forms as the glycoprotein progresses through the Golgi. The sialic acid is present as the terminal sugar on many serum glycoproteins and appears to function in serum longevity and rates of clearance for some serum proteins (Gruenwald et al. 1991). Incorporation of this charged sugar residue into protein glycans has not been demonstrated in plants (Faye et al. 1989). In addition, plants do not phosphorylate high-mannose glycans – in mammals, the mannose-6-phosphate serves as a signal to target soluble glycoproteins to lysosomes. Finally, many complex plant glycans contain either fucose or xylose residues with linkages that do not occur in humans. Plant-synthesized glycoproteins displaying these sugar linkages appear highly immunogenic when injected into mammals (Cumpstey and Faye 1996). Interestingly, an Arabidopsis mutant defective in *N*-acetylglucosaminyl-transferase-I has been identified in which all N-linked glycans are in the high-mannose form (von Schaewen et al. 1993). This report suggests that processing of glycans to complex forms is not critical for plant viability or development (in contrast to animals). Thus, plants can be altered to produce nonimmunogenic glycans. Variations in glycan composition is not unique to plant-based recombinant systems – yeast, baculovirus/insect cell, transgenic animal milk and even mammalian cell cultures often generate glycans that are heterogeneous or differ significantly from the native conformation for particular human proteins (reviewed in Jenkins et al. 1996). It is clear that additional research is required for effective bioproduction of human glycoproteins in plants (discussed further for lysosomal proteins in Sect. 3). Genetic engineering strategies to modify the glycan-processing machinery of plants or *in vitro* enzymatic modification of the purified recombinant protein should enable commercialization of plant-synthesized glycoproteins for pharmaceutical applications.

Because plants are relatively easy to genetically engineer, genetic strategies to specifically alter protein processing by either antisense to block endogenous enzymes or addition of genes encoding novel processing activities are highly feasible. The recent cloning of plant genes encoding enzymes involved in Golgi-localized glycan processing opens up opportunities to modify the complex glycans produced in plants. Processes other than glycosylation can also be modified. We are interested in testing whether plants can be engineered to produce the complex serum proteinases involved in the coagulation-anticoagulation cascade (Cramer et al. 1996a; Wurzburger et al. 1995). Plants are unlikely to perform the highly specialized γ-carboxylation of the amino-terminal glutamates required for bioactivity of several of these enzymes (protein C, thrombin, clotting factors VII, IX and X). We are currently introducing a human cDNA for the vitamin K dependent γ-carboxylase to perform the necessary modifications for this class of proteinase into tobacco (Cramer, Graham, et al. unpublished data). While these experiments are in very early stages, the concept of engineering elite plant lines for specialized protein processing for pharmaceutical bioproduction seems highly feasible.

2.5 Recovery Strategies

To capitalize on the advantages of plant-based systems in upstream production, it is necessary that downstream purification of the recombinant product be accomplished economically. Complex and inefficient purification schemes can contribute significantly to overall costs and result in lower yields so that commercial production is no longer viable. In some cases, such as in the production of industrial enzymes, downstream costs can be reduced or even eliminated when a high degree of product purity is not required. A good example of this is the production of phytase in seeds. The enzyme phytase may be used to enhance the nutritional quality of seed meal by breaking down the phytase present in the meal and thereby increasing the availability of phosphate to monogastric animals. This may be conveniently achieved through expressing the phytase enzymes in seeds and adding milled transgenic seed to a standard feed meal preparation (Pen et al. 1993; Verwoerd and Pen 1996). Unfortunately, this strategy is not applicable to many proteins, particularly pharmaceutical proteins, that require rigorous purification to near-homogeneity. For these products simple and efficient methods of downstream purification must be developed.

2.5.1 Affinity Tag-Based Purification

One approach to the purification of recombinant proteins is through the use of affinity tags. This can be accomplished through the creation of a fusion between the protein of interest and another protein or peptide that exhibits affinity for a specific ligand. The fusion protein is then recovered by binding to the ligand immobilized onto a support matrix. The high selectivity possible with affinity separation often enables a substantial degree of purification to be achieved in a single step. A number of these affinity tags have been developed for use in microbial systems. Different types of ligand pairs have been exploited for this purpose including maltose binding protein–amylose, histidine residues–metal ions, and protein A–IgG. A similar approach may be useful for the purification of recombinant proteins synthesized in plants. The efficacy of this method in plants has been demonstrated in a small scale purification of a human glucocerebrosidase-FLAG epitope fusion produced in tobacco (Cramer et al. 1996b). Here, the fusion protein was recovered using an anti-FLAG antibody affinity matrix and used for bio-

chemical studies on activity and posttranslational modifications. However, because the long-term application is as a replacement enzyme therapeutic for Gaucher patients, the presence of the "nonhuman" residues is undesirable and is not used for scale-up. For some proteins and production strategies, the affinity tag can be proteolytically removed from the fusion protein following purification. However, as with any strategy involving cleavage of fusion proteins, the additional steps required to remove the tag contribute to the downstream purification costs; and there is the potential that the tag could alter folding or processing of the recombinant protein.

2.5.2 Compartmentalization

Another means of simplifying the purification of recombinant proteins is through compartmentalization. This can be achieved using either signal peptides or whole protein fusions to target the protein to a specific cellular location. In this case, purification of the desired protein is facilitated by virtue of its physical separation from other proteins in the cell. Subcellular fractionation is then used to obtain an enriched fraction containing the recombinant protein. A variety of forms of compartmentalization have been recruited for the production of foreign proteins in plants. These include expression on viral particles, extracellular secretion, and targeting to intracellular storage organelles. As noted above, the posttranslational modifications required to produce a functional protein necessarily constrain where that protein can be expressed, as these reactions are, to a large extent, localized to specific subcellular compartments.

A number of plant viruses have been used for the transient expression of foreign proteins in plants (De Zoeten et al. 1989; Gal'a 1993; Usa et al. 1993). To aid in purification, recombinant proteins may be engineered as fusions to viral coat proteins and then separated as mature virus particles. This strategy has recently been used in the production of malarial antigens in tobacco with tobacco mosaic virus (TMV) (Turpen et al. 1995). Nontransgenic plants were inoculated with infectious RNA transcribed from cDNA encoding the genetically engineered fusion protein. Mature viral particles carrying the fusion were later recovered from leaf extracts through differential centrifugation and precipitation. While in this example the purified viral particles were intended for use as a malaria vaccine, it should also be possible to further purify recombinant proteins with this approach by introducing a protease cleavage site into the fusion protein. One possible limitation to this approach may be the size of the foreign protein, as larger proteins may impair viral coat assembly.

Secretion into the extracellular media or periplasmic space has proven to be extremely useful for production and purification of foreign proteins in many yeast and bacterial systems. In addition to providing an enriched fraction of the recombinant product, secretion has also been found to enhance protein stability and facilitate proper folding. Another attractive feature of this approach is that the signal peptide is removed from the recombinant protein in the course of normal processing enabling an authentic protein to be obtained without introducing additional proteolytic digestion steps. In plant cells, secreted proteins are deposited into the apoplastic space. The native signal peptide as well as a signal sequence

from the tobacco pathogenesis-related protein, PR-S, have been used to successfully direct secretion of human serum albumin in potato (Sijmons et al. 1990). Similarly, the potato proteinase inhibitor II protein signal peptide (Hiatasas et al. 1995) has been used to secrete xylanase into the apoplastic space of tobacco plants. While considerable enrichment of the recombinant protein can be achieved with this approach, methods of efficiently recovering proteins from the apoplastic fluid have yet to be developed.

With the appropriate signals or fusions it is also possible to target proteins to the lumen of the ER or vacuole. The human neuropeptide leu-enkephalin has been expressed in seeds of *Arabidopsis thaliana* and *Brassica napus* as an internal fusion between the N- and C-terminal ends of the Arabidopsis 2S albumin protein (Vandekerckhove et al. 1989). The fusion protein was subsequently found to accumulate stably within the protein bodies of these seeds. Purification was accomplished through an initial fractionation in low salt to obtain albumin proteins followed by two proteolytic digestion steps and HPLC separation. One drawback of this strategy is the complexity of the proteolytic cleavage, particularly since carboxypeptidase was required to remove the C-terminal portion of the albumin protein. A failure to precisely control this reaction would result in significant product heterogeneity. It is also possible that folding constraints for protein body packaging might impose limitations on the size of the foreign protein that could be produced as an internal fusion.

Seed oil bodies represent another subcellular compartment available for targeting of recombinant proteins. Localization to oil bodies is achieved through creating a fusion between the desired recombinant protein and oleosin, a protein specifically targeted to these organelles. As described below, oil bodies offer some unique advantages and opportunities for expression and purification of foreign proteins.

2.5.3 Seed Oil Bodies as Purification Tools

Oil bodies are natural subcellular organelles found in all oilseeds where they form the storage site for the primary energy reserve in these seeds, triacylglycerides (TAGs). They are comprised of TAGs surrounded by a half-unit phospholipid membrane into which is embedded a unique type of protein known as oleosin. Oleosins accumulate to high levels in oil seeds comprising between 2% and 10% of the total seed protein in different species. It is believed that the primary function of oleosins is to prevent the coalescence of oil bodies during seed desiccation. In so doing, a larger surface area is available for lipolytic enzymes enabling the rapid mobilization of TAG reserves upon seed germination. Although the precise mechanism of oleosin targeting is not fully understood, it is known that they are synthesized on the ER and that a motif in the central domain is crucial for their subsequent localization to oil bodies (van Rooijen and Moloney 1995a; Abell et al. 1997). The oleosin protein appears to consist of three distinct domains. The N- and C-terminal domains are amphipathic and proteolytic digestion studies strongly suggest that they reside on the outer surface of the oil body (Abell et al.

1997; Hills et al. 1993; Tzen and Huang 1992). The central domain is comprised largely of hydrophobic amino acid residues, and is believed to adopt a hairpin conformation anchoring the protein firmly within the TAG core of the oil body. Comparison of oleosin sequences from different species reveals that the central domain is highly conserved while the N- and C-terminal exhibit considerable sequence variation.

Several features of seed oil bodies lend themselves to the production of foreign proteins. Oleosins tolerate fusion of foreign proteins to either the N- or C-terminal ends without apparent loss of oil body targeting efficiency (Moloney and van Rooijen 1996). Oleosin fusions have been created with a number of different proteins varying in molecular weight from approximately 7–55kDa, all of which are stably accumulated on the surface of oil bodies. In the case of the reporter enzyme β-glucuronidase, it was further shown that enzymatic activity was retained with the oleosin fusion-oil body complex. The oil bodies, together with their complement of oleosin proteins, are remarkably stable both within the seed and following their release by aqueous extraction (van Rooijen and Moloney 1995b). Within the seed the proteins remain undegraded for years without the requirement for elaborate storage conditions. Following their release into aqueous solution, oil bodies are extremely resistant to mechanical disruption and are stable over a wide range in pH and temperature (Kürnel et al. 1996; van Rooijen and Moloney 1995b). Finally, the lower density of oil bodies allows them to be separated from soluble contaminants by flotation centrifugation, enabling simple and rapid purification of recombinant proteins targeted to the oil body surface. Digestion with a site-specific endoproteinase to cleave the oleosin fusion protein, and centrifugation to remove the oil bodies, results in the recovery of a highly enriched fraction of the desired recombinant protein within the aqueous phase. The naturally low hydrolytic environment within the seed, coupled with the rapid removal of soluble-protein contaminants, ensures that little or no degradation of the oil body-associated proteins occurs during processing. As described in Sect. 3.2, the unique properties of oleosins and oil bodies have been exploited by Sem BioSys in the development of a novel plant-based protein production and purification system.

3 Examples of Plant-synthesized Protein Therapeutics: Linking Upstream and Downstream Strategies

In order to "reduce to practice" many of the considerations and strategies described above, two very different examples of plant-based bioproduction of recombinant proteins of commercial value are described below. These examples not only demonstrate the diversity of expression and purification strategies available through plants, but also highlight the constraints on bioproduction strategies imposed by the particular protein target. In both cases, the overall bioproduction strategy has been strongly influenced by commercial and regulatory considerations.

3.1 Production of Human Lysosomal Enzymes in *Nicotiana tabacum*

Cost-effective production of recombinant human proteins for replacement enzyme therapies is likely to have a huge impact on the care and treatment of patients with specific metabolic or genetic disorders. The lysosomal storage disorders represent a large class of these genetic diseases for which the molecular basis of disease has been determined and cDNAs encoding the required enzymes have been cloned (Neufeld 1991). Lysosomes, the animal organelle responsible for the regulated intracellular degradation of macromolecules, contain multiple hydrolases including proteases, nucleases, glycosidases, lipases, phosphatases, phospholipases, and sulfatases (Darnell et al. 1986). Deficiency in specific lysosomal hydrolases can lead to toxic accumulation of the undegraded substrate and a variety of clinical manifestations. Tay-Sachs disease is perhaps the most familiar lysosomal storage disease, involving deficiencies in α-hexosaminidase that lead to accumulation of ganglioside G_{M2} in the membrane of brain cells (Neufeld 1991). The mucopolysaccharidoses (MPSs) are a group of lysosomal storage diseases caused by deficiencies of one or more of the ten lysosomal enzymes required for the degradation of sulfated glycosaminoglycans (reviewed in Neufeld and Muenzer 1995). Lysosomal accumulation of undegraded glycans leads to the malfunction of affected cells/organs which compromises the growth and development of the individual and may, in severe cases, lead to premature death. Replacement enzyme therapy appears promising based on human cell- and animal models, but drug development is hampered by the small patient pool and limitation in current technologies for cost-effective bioproduction. The industry paradigm for human replacement enzyme therapy is the glycoprotein product Ceredase (Genzyme, Cambridge, MA) for the treatment of Gaucher disease. This lysosomal storage disorder affects 10,000–20,000 individuals in the United States (NIH Technology Assessment Panel on Gaucher Disease 1996) and is caused by defects in glucocerebrosidase, an acid β-glucosidase required for complex lipid degradation. Routine administration (generally every 2 weeks) of placental-derived enzyme has revolutionized the treatment of the disease and the quality of life of Gaucher patients. However, the high drug costs associated with purification of glucocerebrosidase from human placentae or, more recently, with bioproduction of recombinant enzyme in Chinese hamster ovary (CHO) cells, make it one of the world's most expensive drugs. Although the production of lysosomal enzymes in plants is challenging (Chakrabarty et al. 1996b), CropTech has selected several lysosomal enzymes among its initial targets for bioproduction based on (a) the ability of plants to address critical cost, safety and supply issues for replacement enzymes, (b) the extreme medical need, and (c) the potential for Orphan Drug status to facilitate progress toward clinical trials and commercialization.

The first lysosomal enzyme produced in transgenic plants was glucocerebrosidase (EC3.2.1.45) as a potential alternative replacement therapy for Gaucher disease (Chakrabarty et al. 1996a,b). Placental glucocerebrosidase that has been enzymatically modified to generate mannose-terminated glycans is highly effective in

reversing the symptoms of the disease (Brady et al. 1974). For commercial production of glucocerebrosidase (Ceredase, Genzyme), it requires between 400 and 2000 placentae to supply a standard dose – a major factor in the extreme cost to patients ($100,000–400,000 annually; NIH Technology Assessment Panel on Gaucher Disease 1996). A CHO-synthesized recombinant form (Cerezyme, Genzyme) has been approved by the FDA, but no significant reduction in cost is anticipated. As a consequence, the success of treatment of Gaucher's disease in the United States remains limited by the cost and supply of the drug. For these reasons, it is a promising candidate for production in a plant-based system. The successful synthesis of enzymatically active human glucocerebrosidase in transgenic tobacco plants was previously described (Cramer et al. 1996a,b). Briefly, the entire coding region for human glucocerebrosidase was modified to encode a FLAG (International Biotechnologies) epitope tag for subsequent detection and purification, fused to the inducible MeGA[TM] promoter (CropTech Corporation), and introduced into tobacco. Of 49 transgenic plants analyzed, five plants had expression levels between 1%–10% total soluble leaf protein after 12h of postharvest induction. Based on biochemical kinetic studies, affinity binding characteristics, and conduritol B inhibition, the tobacco-synthesized enzyme appears fully active and comparable to human- or CHO-derived glucocerebrosidase. These studies form the basis of current and future efforts in the commercial production of glucocerebrosidase in transgenic plants for Gaucher enzyme replacement therapy. Studies to (a) identify high-expressing transgenic lines containing new glucocerebrosidase constructs that lack the nonhuman FLAG epitope and (b) test strategies to address glycan modification are now underway prior to scale-up of glucocerebrosidase production and purification technologies.

CropTech researchers have also synthesized a second lysosomal enzyme, α-L-iduronidase (IDUA, EC3.2.1.76), in transgenic tobacco (Jenkins, Weissenborn, Bennett, and Oishi, unpublished results). IDUA is a potential replacement therapeutic for Hurler syndrome and Hurler/Scheie syndrome, the most common MPS representing 1/100,000–1/150,000 births. Although the concept of enzyme replacement for Hurler syndrome was first investigated in the 1970s (DiFerrante and Nichols 1974; Neufeld and Muenzer 1989), the development of IDUA as a drug has not progressed rapidly because of the lack of an effective production system. Recombinant enzyme sufficient for initial testing in Hurler-canine and feline models has been produced using a CHO-based production system (Kakkis et al. 1994), but progress toward human trials is limited by protein availability. As a consequence, the successful plant-based production of IDUA has the potential to directly impact the speed of development of IDUA as a enzyme replacement therapy for Hurler and Hurler/Scheie syndromes. In humans, the lysosomal IDUA from liver is a soluble glycoprotein of 60–82kDa reflecting heterogeneity in the glycan composition. There are six potential N-linked glycosylation sites, some of which are modified to mannose-6-phosphate forms (generally sites 3 and 6) or to complex glycan forms. At all sites there is a high degree of microheterogeneity in glycans (Zhao et al. 1996). The sequence of the complete cDNA for human IDUA has been reported (Moskowitz et al. 1992; Scott et al. 1991) and encodes a protein

of 653 amino acids (pre-IDUA) with a signal peptidase cleavage site at amino acid 27. The cDNA for IDUA has been expressed in Cos-1 and CHO cells (Kakkis et al. 1994; Scott et al. 1991) and recombinant IDUA has been purified and shown to be biologically active in dogs deficient for this enzyme (Kakkis et al. 1996).

As an initial test of the feasibility of commercial production of human IDUA in plants, researchers at CropTech Corporation engineered tobacco plants for postharvest production of human IDUA. The coding region from the human IDUA cDNA (Kakkis et al. 1994) was placed downstream of the inducible MeGA[TM] promoter and inserted into a plant transformation vector for *Agrobacterium*-mediated transformation (Horsch et al. 1985). Transgenic tobacco plants containing one to three copies of the MeGA promoter-IDUA construct were tested for IDUA expression. In noninduced maturing transgenic leaf material, no IDUA mRNA was detected. Following 8–24h of induction, abundant IDUA transcripts were detected as well as novel proteins that cross-reacted with anti-IDUA antisera. Tobacco contains no endogenous IDUA activity (in contrast to the acid β-glucosidases which are quite abundant in tobacco). Analyses of IDUA recovered from induced tobacco leaf material indicated that the IDUA protein is enzymatically active. Because human IDUA requires glycosylation for activity, detection of enzyme activity in tobacco samples (as well as its electrophoretic mobility and ConA binding) indicates that the human IDUA signal peptide correctly targets the protein to the plant endomembrane system for glycosylation. The majority of the IDUA appears to be secreted, the default pathway for the plant endomembrane system (Denecke et al. 1990). Although IDUA yields from the first IDUA-transgenic plants analyzed are lower than those seen for glucocerebrosidase-expressing plants, demonstrations of enzymatic activity of the tobacco-synthesized IDUA glycoprotein and development of novel IDUA recovery methods strongly support the use of transgenic tobacco for human IDUA production.

Both glucocerebrosidase and IDUA are glycoproteins and thus pose a particular challenge for production in plants as well as other recombinant expression systems (Jenkins et al. 1996). For soluble lysosomal enzymes such as IDUA, the signal for lysosomal sorting is the mannose-6 phosphate residues present in their N-linked glycans. Mannose-6-phosphate receptors are present on the plasmamembrane as well as lysosomal membranes of many mammalian cell types and thus direct uptake and lysosomal delivery of exogenously supplied IDUA (Kakkis et al. 1996). Plants do not phosphorylate their glycans and glycan-based signals do not appear to function in vacuolar targeting (Faye et al. 1989; Chrispeels and Faye 1996). It is likely that some, if not all, of the glycans on tobacco-synthesized IDUA are in the complex form and thus likely to be immunogenic (see Sect. 2.4) and ineffective in directing the required cell-specificity for uptake and lysosomal delivery. Engineering plants to synthesize mannose-6-phosphate-modified glycans is currently not feasible – the two required enzymes have not been well characterized. However, alternative strategies that address both the delivery and immunogenicity are suggested by the currently effective lysosomal replacement therapeutic, Ceredase. Glucocerebrosidase is a membrane-associated protein that is targeted to lysosomes by a mannose-6-phosphate-independent route. The N-linked glycans

present on the placental enzyme are biantenulary structures having terminal sialic acid residues. In order to direct effective delivery to lysosomes of the affected cells in Gaucher patients (primarily cells of the macrophage/monocyte lineage), sequential enzymatic digestion is used to remove the terminal sugars and expose the mannose core (Barton et al. 1991). This mannose-terminated form is targeted to the correct cell and organellar location to effect glucocerebroside degradation and symptom reduction (Grabowski et al. 1995; Barton et al. 1991). Complex plant glycans are naturally mannose terminated (Chrispeels and Faye 1996). Enzymatic removal of the immunogenic fucose and xylose residues should yield glycans of similar pharmacokinetics as Ceredase.

3.2 Production of Hirudin in *Brassica napus*

To evaluate the potential of Sem BioSys' oleosin partitioning technology, the model therapeutic protein hirudin was selected. Hirudin is a naturally occurring anticoagulant protein produced in the salivary glands of medicinal leeches (*Hirudo medicinalis*) and secreted to facilitate feeding. Since its discovery almost 50 years ago, it has been extensively studied. Hirudin possesses a number of desirable properties which advocate its use as a therapeutic pharmaceutical. It is an extremely specific and potent inhibitor of thrombin, the last enzyme in the blood coagulation cascade, having a KI of 2 fM (Braun et al. 1988). It is also rapidly cleared from the body, exhibits low toxicity (500,000 U/kg body weight in rats) (Markwardt et al. 1982) and, probably as a consequence of the coevolution of leeches and mammals, has relatively low immunogenicity (Klocking 1991). The protein has also been well characterized with respect to its structure and mechanism of binding to thrombin (Rydel et al. 1990). A small number of closely related isoforms of hirudin have been isolated all of which show strict conservation for six cysteine residues (Stone and Maraganore 1993). These residues participate in the formation of three disulfide bridges whose precise pairing is necessary for protein activity (Chatrenet and Chang 1992, 1993). Although the native protein is sulfated at the Tyr-63 position, recombinant nonsulfated hirudin exhibits significant activity (Stone and Maraganore 1993). It folds spontaneously *in vitro* and functional hirudin has been produced previously in both bacterial (Harvey et al. 1986; Bindar et al. 1990) and yeast (Loison et al. 1988; Lemoan et al. 1993) systems. However, the quantities of hirudin required, were it to fully replace presently used anticoagulants such as heparin, are estimated to be on the order of hundreds to thousands of kilograms of protein annually. For this reason, hirudin is an excellent candidate for production with a high capacity plant-based system.

The common oilseed rape species, *Brassica napus*, was selected as the vehicle for production of seed-derived hirudin. After tobacco, the *Brassica* species are among those most easily transformed with *Agrobacterium*. Cells in the ends of cotyledonary petioles cut from young seedlings are readily infected with the bacterium. Formation of callus, regeneration to plants, and selection of transformants are all very efficient. In *B. napus*, transformation efficiencies approaching 55% of

the original explants can be obtained. The time-line for development of a transgenic plant is also relatively short, in the range of approximately 4–6 months from transformation to collection of first generation transformed seed. Another attractive feature is the availability of a haploid production system from microspore-derived embryos, facilitating the creation of homozygous lines. As an oilseed crop, considerable biomass is concentrated within the seed. Seed production in *B. napus* is between 1 and 2 tons per hectare at a cost of approximately (United States) $300/ton. Protein content in these seeds represents in excess of 20% of the total seed weight, approximately 9% of which is oleosin.

The production and analysis of transgenic plants expressing an oleosin-hirudin fusion has been reported previously (Parmenter et al. 1995). Briefly, a synthetic sequence encoding the hirudin variant 2 (HV2) isoform was fused to the 3' end of an Arabidopsis 19kDa oleosin gene with the two coding regions separated by a sequence encoding the four amino acid recognition site for the protease, factor Xa. Following *Agrobacterium*-mediated transformation, putative transgenics were selected and expression of the oleosin-hirudin fusion confirmed by northern analysis. Immunoblotting with anti-hirudin antibodies demonstrated that the oleosin-hirudin fusion protein was correctly targeted and accumulated on oil bodies of transgenic seed. Oil bodies were separated and washed to remove contaminating soluble proteins through flotation-centrifugation. After digestion with factor Xa and a final round of flotation-centrifugation to remove oil bodies, hirudin was recovered in the aqueous fraction. Formation of a functional protein was confirmed by an *in vitro* thrombin inhibition assay. Comparison of protein contents in whole seed extracts and in the soluble fraction obtained after flotation-centrifugation indicated that the majority of seed protein had been removed. The enrichment obtained with this procedure demonstrates the utility of oil body compartmentalization for purification of recombinant proteins. Further purification of the recombinant hirudin to near-homogeneity was achieved through anion exchange and reverse phase chromatography. Values obtained for the specific activity of *B. napus*-derived hirudin are equivalent to those reported for recombinant hirudin produced in yeast systems (Loison et al. 1988).

3.2.1 Prospects of Oleosin-Partitioning Technology

The potential for commercial application of oleosin partitioning technology can be evaluated by examining the system with reference to certain key production parameters namely, production capacity, authenticity/functionality of product, downstream purification costs, and process scaleability. We have estimated the level of expression of the oleosin-hirudin fusion protein in our transgenic seed to be approximately 10% of that of the endogenous oleosin (Parmenter et al. 1995). Based on this estimate, hirudin would represent approximately 0.3% of the total seed protein. While encouraging, this level is still somewhat lower than would be desired for a commercial production system. To increase expression levels, we are currently testing a number of strong seed-specific promoters other than oleosin in our fusion constructs. An increase in the expression of recombinant protein to the relatively

modest level of 1% of seed protein would result in a system capacity of approximately 2kg of product per ton of seed. When coupled with low production costs and cost-effective purification, this level is within the range required for commercial viability.

The downstream purification of proteins synthesized as oleosin fusions is greatly simplified by the oil body separation process. However, in order for this process to be cost-effective, the fusion protein cleavage step must be both efficient and economical. While useful for demonstration purposes, the factor Xa used in our initial hirudin studies fails to meet these requirements. The enzyme is expensive, gave incomplete cleavage, and represented a contaminant which had to be removed in subsequent purification steps. To address this problem we are presently expressing proteases as oleosin fusions immobilized on the surface of oil bodies. This will enable both economical production of the protease and easy removal following fusion protein cleavage through the existing oil body separation process. A number of suitable candidate proteases have been identified and are currently being tested.

The importance of process scale-up in determining economic feasibility is often overlooked in the initial research and development phase of a new technology. Procedures that work well for typical laboratory scale experiments cannot always be directly scaled up or easily adapted to existing industrial processes. In the case of oleosin partitioning technology, we have developed and tested methods using industrial equipment for the large scale preparation of oil bodies. The results from these tests indicate that the process can be easily scaled up to meet commercial production requirements.

The recovery of active products such as hirudin and β-glucuronidase from oleosin fusions demonstrates that functional proteins can be produced using oleosin partitioning technology. However, the fact that oleosins are not exposed to the lumen of the ER either during synthesis or subsequent targeting to oil bodies, limits the range of different products that can be produced through oleosin partitioning. Proteins requiring glycosylation or other forms of posttranslational modification associated with passage through the endomembrane system would not be properly processed as oleosin fusions. Nevertheless, a large number of proteins are still amenable to production using this technology, in addition to therapeutic proteins, the list includes many food and industrial enzymes. Some of these products are presently under development. Additionally, the ability to produce functional proteins on the surface of oil bodies offers exciting new possibilities for the production of immobilized protein matrices (Kretzmer et al. 1996). With continued development in each of the areas mentioned above, prospects for the successful commercialization of oleosin partitioning technology appear very promising.

4 Summary

We have described two very different and innovative plant-based production systems – postharvest production and recovery of recombinant product from tobacco leaves using an inducible promoter and oleosin-mediated recovery of recombinant product from oilseeds using a seed-specific promoter. Both base technologies are broadly applicable to numerous classes of pharmaceutical and industrial proteins. As with any emerging technology, the key to success may lie in identifying those products and applications that would most benefit from the unique advantages offered by each system. The postharvest tobacco leaf system appears effective for proteins requiring complex posttranslational processing and endomembrane targeting. Because of the remarkable fecundity and biomass production capacity of tobacco, biomass scale-up is very rapid and production costs are low. Clearly the development of equally cost-effective extraction and purification technologies will be critical for full realization of the commercial opportunities afforded by transgenic plant-based bioproduction. The recovery of protein from tobacco leaves or oleosin-partitioned proteins by oil-body separation represent significant breakthroughs for cost-effective commercialization strategies. Additional low-cost, high-affinity separation technologies need to be developed for effective scale-up purification of plant-synthesized recombinant proteins. Clearly successful commercialization of plant-synthesized biopharmaceuticals must effectively link upstream strategies involving gene and protein design with downstream strategies for reproducible GMP-level recovery of bioactive recombinant protein. Both the tobacco and oilseed systems are uniquely designed to address issues of biomass storage, product recovery, quality assurance, and regulatory scrutiny in addition to issues of transgene expression and protein processing.

References

Abell BM, Holbrook LA, Abenes M, Murphy DJ, Hills MJ, Moloney MM (1997) Role of the proline knot motif in oleosin topogenesis and oleosin topology-independent oil body targeting. Plant Cell 9:1481–1493

Barton NW, Brady RO, Dambrosia JM, Bianchine AM, Doppelt SH, Hill SC, Mankin HJ, Murray GJ, Parker RI, Argoff CE, Grewal RP, Yu K-T (1991) Replacement therapy of inherited enzyme deficiency: macrophage-targeted glucocerebrosidase for Gaucher's disease. N Engl J Med 324:1464

Bayan N, Vogel R, Kellis K, Lauga J (1990) Synthesis and secretion of hirudin by *Streptomyces lividans*. Appl Microbiol Biotech 34:203–207

Beachy PN, Ling B, Chen M-H (1989) The CaMV 35S enhancer contains at least two domains which can confer different developmental and tissue-specific expression patterns. EMBO J 8:2195–2202

Bosch D, Smid J, Krebber S (1994) A trout growth hormone is expressed, correctly folded and partially glycosylated in the leaves but not the seed of transgenic plants. Transgenic Res 3:304–310

Brady RO, Pentchev PG, Gal AE, Hibbert SR, Dekaban AS (1974) Replacement therapy for inherited enzyme deficiency. Use of purified glucocerebrosidase in Gaucher's disease. N Engl J Med 291:989–993

Braun PJ, Dennis S, Hofsteenge J, Stone SR (1988) Use of site-directed mutagenesis to investigate the basis for the specificity of hirudin. Biochemistry 27:6517–6522

Carrington JC, Freed DD (1990) Cap-independent enhancement of translation by a plant potyvirus 5′ nontranslated region. J Virol 64:1590–1597

Chatrenet B, Chang J (1993) The folding of hirudin adopts a mechanism of trial and error. J Biol Chem 267:3038–3043

Chatrenet B, Chang J (1993) The disulfide folding pathway of hirudin elucidated by stop/go folding experiments. J Biol Chem 268:20965–20996

Chan AM, Cemal L (1996) Gene transcription. In: Owen MRL, Pen J (eds) Transgenic plants: a production system for industrial and pharmaceutical proteins. Wiley, Chichester, pp 27–48
Chong DKX, Roberts W, Arakawa T, Illes K, Bagi G, Slattery CW, Langridge WHR (1997) Expression of the human milk protein b-casein in transgenic potato plants. Transgenic Res 6:289–296
Chrispeels MJ, Faye L (1996) The production of recombinant glycoproteins with defined non-immunogenic glycans. In: Owen MRL, Pen J (eds) Transgenic plants: a production system for industrial and pharmaceutical proteins. Wiley, Chichester, pp 99–114
Cramer CL, Weissenborn DL, Oishi KK, Grabau EA, Bennett S, Ponce E, Grabowski GA, Radin DN (1996a) Bioproduction of human enzymes in transgenic tobacco. In: Collins GB, Shepherd RJ (eds) Engineering Plants for Commercial Products and Applications. New York Academy of Sciences, New York, pp 62–71
Cramer CL, Weissenborn DL, Oishi KK, Radin DN (1996b) High-level production of enzymatically active human lysosomal proteins in transgenic plants. In: Owen MRL, Pen J (eds) Transgenic plants: a production system for industrial and pharmaceutical proteins. Wiley, Chichester, pp 299–310
Cramer CL, Weissenborn DL (1997) HMG2 promoter expression system and post-harvest production of gene products in plants and plant cell cultures. US Patent 5,670,349
Darnell J, Lodish H, Baltimore D (1986) Molecular cell biology. Scientific American, New York
Denecke J, Botterman J, Deblaere R (1990) Protein secretion in plant cells can occur via a default pathway. Plant Cell 2:51–59
de Zoeten GA, Penswick JR, Horisberger MA, Ahl P, Schultze M, Hohn T (1989) The expression, localization and effect of a human interferon in plants. Virology 172:213–222
Dieryck W, Pagnier J, Poyart C, Marden MC, Gruber V, Bournat P, Baudino S, Merot B (1997) Human haemoglobin from transgenic tobacco. Nature 386:29–30
DiFerrante NM, Nichols BL Jr (1974) Comments on the plasma treatment of the mucopolysaccharidoses. Birth Defects 10:234–238
Fang R-X, Nagy F, Sivasubramaniam S, Chua N-H (1989) Multiple cis-regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. Plant Cell 1:141–150
Faye L, Johnson KD, Sturm A, Chrispeels MJ (1989) Structure, biosynthesis, and function of asparagine-linked glycans on plant glycoproteins. Plant Physiol 75:309–314
Firek S, Draper J, Owen MRL, Gandecha A, Cockburn B, Whitelam GC (1993) Secretion of a functional single-chain Fv protein in transgenic tobacco plants and cell suspension cultures. Plant Mol Biol 23:861–870
Ganz PR, Dudani AK, Tackaberry ES, Sardana R, Sauder C, Cheng X, Altosaar I (1996) Expression of human blood proteins in transgenic plants: the cytokine GM-CSF as a model protein. In: Owen MRL, Pen J (eds) Transgenic plants: a production system for industrial and pharmaceutical proteins. Wiley, Chichester, pp 281–297
Grabowski GA, Barton NW, Pastores G, Dambrosia JM, Banerjee TK, McKee MA, Parker C, Schiffmann R, Hill SC, Brady RO (1995) Enzyme therapy in type 1 Gaucher disease: comparative efficacy of mannose-terminated glucocerebrosidase from natural and recombinant sources. Ann Intern Med 122:33–39
Grill LK (1993) Tobacco mosaic virus as a gene expression vector. Agro-Food-Industry Hi-Tech November/December: 20–23
Grill LK (1997) Viral-vectored, large scale production of drugs and pharmaceuticals in plants. Presentation at IBC's 3rd Annual International Symposium on Producing the Next Generation of Therapeutics: Exploiting Transgenic Technologies, West Palm Beach, 5–6 Feb
Grinnell BW, Walls JD, Gerlitz B (1991) Glycosylation of human protein C affects its secretion, processing, functional activities and activation by thrombin. J Biol Chem 266:9778–9785
Haq TA, Mason HS, Clements JD, Arntzen CJ (1995) Oral immunization with a recombinant bacterial antigen produced in transgenic plants. Science 268:714–716
Harvey RP, Degryse E, Stefani L, Schamber F, Cazenave JP, Courtney M, Tolstoshev P, Lecocq JP (1986) Cloning and expression of cDNA coding for anticoagulant hirudin from blood-sucking leech, *Hirudo medicinalis*. Proc Natl Acad Sci USA 83:1084–1088
Hein MB, Yeo T-C, Wang F, Sturtevant A (1996) Expression of cholera toxin subunits in plants. In: Collins GB, Shepherd RJ (eds) Engineering plants for commercial products and applications. New York Academy of Sciences, New York, pp 50–56
Herbers K, Wilke I, Sonnewald U (1995) A thermostable xylanase from *Clostridium thermocellum* expressed at high levels in the apoplast of transgenic tobacco has no detrimental effects and is easily purified. Biotechnology 13:63–66

Higo K, Saito Y, Higo H (1993) Expression of a chemically synthesized gene for human epidermal growth factor under the control of cauliflower mosaic virus 35S promoter in transgenic tobacco. Biosci Biotechnol Biochem 57:1477–1481
Hills MJ, Watson MD, Murphy DJ (1993) Targeting of oleosins to the oil bodies of oilseed rape. Planta 189:24
Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT (1985) A simple and general method for transferring genes into plants. Science 227:1229–1231
Jenkins N, Parekh RB, James DC (1996) Getting the glycosylation right: implication for the biotechnology industry. Nat Biotechnol 14:975–981
Kay R, Chan A, Dal M, McPherson J (1987) Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. Science 236:1299–1302
Kakkis ED, Matynia A, Jonas AJ, Neufeld EF (1994) Overexpression of the human lysosomal enzymes α-L-iduronidase in Chinese hamster ovary cells. Prot Express Pur 5:225–232
Kakkis ED, McEntee MF, Schmidtchen A, Neufeld EF, Ward DA, Gompf RE, Kania S, Bedolla C, Chien SL, Shull RM (1996) Long-term and high-dose trials of enzyme replacement therapy in the canine model of mucopolysaccharidosis I. Biochem Mol Med 58:156–167
Klocking H (1991) Toxicology of hirudin. Sem Thromb Hemostas 17:126–129
Kühnel B, Holbrook LA, Moloney MM and van Rooijen GJH (1996) Oil bodies of transgenic *Brassica napus* as a source of immobilized β-glucuronidase. J Am Oil Chem Soc 73:1533–1538
Kumagai MH, Turpen TH, Weinzettl N, Della-Cioppa G, Turpen AM, Donson J, Hilf ME, Grantham GL, Dawson WO, Chow TP, Piatak MJ, Grill LK (1993) Rapid high-level expression of biologically active α-trichosanthin in transfected plants by an RNA viral vector. Proc Natl Acad Sci USA 90:427–430
Lehman ED, Joyce JG, Bailey FJ, Markus HZ, Schultz LD, Dunwiddie CT, Jacobson MA, Miller WJ (1993) Expression, purification and characterization of milligram amounts of a recombinant hybrid HV1-HV2 hirudin variant expressed in *Saccharomyces cerevisiae*. Protein Expr Pur 4:247–255
Lindsey K (1996) Plant transformation systems. In: Owen MRL, Pen J (eds) Transgenic plants: a production system for industrial and pharmaceutical proteins. Wiley, Chichester, pp 5–26
Loison G, Findeli A, Bernard S, Nguyen-Juilleret M, Marquet M, Riehl-Bellon N, Carvallo D, Guerra-Santos L, Brown SW, Courtney M, Roitsch C, Lemoine Y (1988) Expression and secretion in *S. cerevisiae* of biologically active leech hirudin. Biotechnology 6:72–77
Lui DTY (1992) Glycoprotein pharmaceuticals – scientific and regulatory considerations and the United States Orphan Drug Act. Trends Biotechnol 10:114–120
Ma JK-C, Hiatt A, Hein MB, Vine ND, Wang F, Stabila P, van Dolleweerd C, Mostov K, Lehner T (1995) Generation and assembly of secretory antibodies in plants. Science 268:716–719
Ma S, Zhao D, Yin A, Mukherjee R, Singh B, Qin H, Stiller CR, Jevnikar AM (1997) Transgenic plants expressing autoantigens fed to mice induce oral immune tolerance. Nat Med 3:793–796
Markwardt F, Hauptmann J, Nowak G, Klebn C, Walsmann P (1982) Pharmacological studies on the antithrombotic action of hirudin in experimental animals. Thromb Haemostasis 47:226–229
Mason HS, Lam D, Arntzen CJ (1992) Expression of hepatitis B surface antigen in transgenic plants. Proc Natl Acad Sci USA 89:11745–11749
Matsumoto S, Ikura K, Ueda M, Sasaki R (1995) Characterization of a human glycoprotein (erythropoietin) produced in cultured tobacco cells. Plant Mol Biol 27:1163–1172
Miele L (1997) Plants as bioreactors for biopharmaceuticals: regulatory considerations. Trends Biotechnol 15:45–50
Moloney MM and van Rooijen GJH (1996) Recombinant proteins via oleosin partitioning. Inform 7:107–113
Moskowitz SM, Dlott B, Chuang PD, Neufeld EF (1992) Cloning and expression of a cDNA encoding the human lysosomal enzyme α-L-iduronidase. FASEB J 6:A77 (abstract)
NIH Technology Assessment Panel on Gaucher Disease (1996) Gaucher disease: current issues in diagnosis and treatment. JAMA 275:548–553
Neufeld EF (1991) Lysosomal storage diseases. Annu Rev Biochem 60:257–280
Neufeld EF, Muenzer J (1989) The mucopolysaccharidoses. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The metabolic basis of inherited disease 6th edn. McGraw-Hill, New York, pp 1565–1587
Neufeld EF, Muenzer J (1995) The mucopolysaccharidoses. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The metabolic and molecular bases of inherited disease 7th edn. McGraw-Hill, New York, pp 2465–2494
Owen MRL, Pen J (1996) Transgenic plants: a production system for industrial and pharmaceutical proteins. Wiley, Chichester

Parmenter DL, Boothe JG, Van Rooijen GJH, Yeung EC, Moloney MM (1995) Production of biologically active hirudin in plant seeds using oleosin partitioning. Plant Mol Biol 29:1167–1180
Pen J (1996) Comparison of host systems for the production of recombinant proteins. In: Owen MRL, Pen J (eds) A production system for industrial and pharmaceutical proteins. Wiley, Chichester, pp 149–168
Pen J, Verwoerd TC, van Paridon PA, Beudeker RF, van den Elzen PJM, Geerse K, van der Klis JD, Versteegh HAJ, van Ooyen AJJ, Hoekema A (1993) Phytase-containing transgenic seed as a novel feed additive for improved phosphorus utilization. Biotechnology 11:811–814
Roberts GW, James S (1996) Prion diseases: transmission from mad cows? Curr Biol 6:1247–1249
Rydel TJ, Ravichandran KG, Tulinsky A, Bode W, Huber R, Roitsch C, Fenton JW II (1990) The structure of a complex of recombinant hirudin and human α-thrombin. Science 249:277
Scott HS, Nelson PV, Hopwood JJ, Morris CP (1991) Human α-L-iduronidase: cDNA isolation and expression. Proc Natl Acad Sci USA 88:9695–9699
Sijmons PC, Dekker BMM, Schrammeijer B, Verwoerd TC, van den Elzen PJM, Hoekema A (1990) Production of correctly processed human serum albumin in transgenic plants. Biotechnology 8:217–221
Stone SR, Maraganore JM (1993) Hirudin and hirudin-based peptides. Methods Enzymol 223:312–336
Taylor CB (1997) Comprehending cosuppression. Plant Cell 9:1245–1249
Turpen TH, Reinl SJ, Charoenvit Y, Hoffman SL, Fallarme V, Grill LK (1995) Malarial epitopes expressed on the surface of recombinant tobacco mosaic virus. Biotechnology 13:53–57
Tzen JTC, Huang AHC (1992) Surface structure and properties of plant seed oil bodies. J Cell Biol 117:327–335
Usha R, Rohll JB, Spall VE, Shanks M, Maule AJ, Johnson JE, Lomonossoff GP (1993) Expression of an animal virus antigenic site on the surface of a plant virus particle. Virology 197:366–374
van Rooijen GJH, Moloney MM (1995a) Structural requirements of oleosin domains for the subcellular targeting to the oil body. Plant Physiol 109:1353–1361
van Rooijen GJH, Moloney MM (1995b) Plant seed oil-bodies as carriers for foreign proteins. Bio-technology 13:72–77
Vandekerckhove J, Van Damme J, Van Lijsebettens M, Botterman J, De Block M, Vandewiele M, De Clercq A, Leemans J, Van Montagu M, Krebbers E (1989) Enkephalins produced in transgenic plants using modified 2S seed storage proteins. Biotechnology 7:929–932
Varki A (1993) Biological roles of oligosaccharides: all of the theories are correct. Glycobiology 3:97–130
Vaughan P (1996) Creutzfeldt-Jakob disease-latest unknown in struggle to restore faith in the blood supply. Can Med Assoc J 155:563–568
Verwoerd TC, Pen J (1996) Phytase produced in transgenic plants for use as a novel feed additive. In: Transgenic plants, a production system for industrial and pharmaceutical proteins. Wiley, Chichester, pp 213–223
von Schaewen A, Sturm A, O'Neill J, Chrispeels MJ (1993) Isolation of a mutant *Arabidopsis* plant that lacks N-acetyl glucosaminyl transferase I and is unable to synthesize Golgi-modified complex N-linked glycans. Plant Physiol 102:1109–1118
Weissenborn DL, Gruber HA, Pedersen K, Oishi KK, Bennett S, Velander W, Cramer CL (1995) Production of human protein C in transgenic tobacco. International Conference on Engineering Plants for Commercial Products and Applications, Lexington KY, October, 1995 (abstract 50)
Zhao HG, Li HH, Bach G, Schmidtchen A, Neufeld EF (1996) The molecular basis of San Filippo syndrome type B. Proc Natl Acad Sci USA 12:6101–6105
Zhu Z, Hughes KW, Huang L, Sun B, Liu C, Li Y, Hou Y, Li X (1994) Expression of human α-interferon cDNA in transgenic rice plants. Plant Cell Tissue Org Cult 36:197–204